JOURNAL OF COMPLEXITY **13,** 480–488 (1997)
ARTICLE NO. CM970457

Polya's Characterization Theorem for Complex Random Variables*

Nicholas N. Vakhania†

Institute of Computational Mathematics, Academy of Sciences of Georgia, Tbilisi 93, Georgia

Received April 1, 1997

A complex random variable can be Gaussian in either the narrow or the wide sense. It is observed that Gaussian random variables in the wide sense do not have the 2-stability property (which is well known for the real case), while in the narrow definition they do possess it. Moreover, it is proved that this property characterizes the class of complex Gaussian random variables in the narrow sense; no other complex random variable enjoys it. © 1997 Academic Press

1. INTRODUCTION

This paper is a natural extension of a previous investigation for the real case setting (see [1]), which was motivated by the following question that was posed to the author by J. F. Traub and H. Woźniakowski: Is the average case norm of a linear unbounded densely defined operator A in a real Banach space X finite with respect to a Gaussian measure μ in X? In other words, we want to know whether the integral

$$\int_X \|Af\|^2 \mu(df)$$

exists and is finite. Of course, we should assume that the linear manifold D_A (the domain of A) is a measurable set and, moreover, the equality $\mu(D_A) = 1$ holds.

This question is of interest in information-based complexity theory, as it is directly tied with the average case complexity of linear ill-posed problems (see [2]).

*This work was partially supported by the Grant N1.16 of the Georgian Academy of Sciences.
†E-mail: vakhania@acnet.ge.

0885-064X/97 $25.00

The answer to this question is affirmative, and this can be proved by two different approaches (see [3]). One of these two proofs is based on the following characterization property of real Gaussian random variables, known as *Polya's characterization theorem*; see [6].

THEOREM 1 [6]. *Let ξ be a real random variable, let n be a natural number, let $\xi_1, \xi_2, \dots, \xi_n$ be mutually independent copies of ξ, and let $a_1, a_2, \dots, a_n$ be nonzero real numbers such that $a_1^2 + a_2^2 + \cdots + a_n^2 = 1$. If the variable $a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$ has the same distribution as ξ then ξ is Gaussian.*

In [1], we gave an elementary and simple proof of the theorem in this general setting (Polya's original proof assumed existence of the second moment for ξ). Here we will show that with a natural necessary modification, the analogous result is valid for complex random variables as well. The scheme of the proof is similar to that in the real case. It is based on three elementary auxiliary statements, which are given below as Lemmas 1–3. Note that the real-variable versions of Lemmas 2 and 3 have been used in the real case. However, the definition of complex Gaussian random variables needs to be properly specified, and so we start with comments on the two possible definitions (see also [4] for more details).

A *complex random variable* $\xi = \xi' + i\xi''$ is, in fact, a pair of real random variables ξ' and ξ'', and it can therefore be considered as a two-dimensional real random vector. Accordingly, it can be called *Gaussian* if the pair (ξ', ξ'') is a Gaussian random vector in $\mathbb{R}^2$. This is the definition in the *wide sense*. The other possibility, that is commonly accepted, is the definition in the *narrow sense*, which assumes additionally that ξ' and ξ'' are independent and have the same variances.

The usual motivation for the narrow definition comes from the form of the characteristic function of a centered Gaussian random variable. For the real case, this is given as

$$\exp\{-\tfrac{1}{2}\, t^2 E\xi^2\} \qquad \forall\, t \in \mathbb{R}^1,$$

with E denoting expectation (mean value). Now recall that the characteristic function χ_ξ of a complex random variable ξ is defined as

$$\chi_\xi(z) = E\ \exp\{i\, \mathrm{Re}\, \xi\bar{z}\} \qquad \forall\, z \in \mathbb{C}^1, \tag{1}$$

where $\mathbb{C}^1$ is the complex plane, $\bar{z}$ denotes the complex conjugate to z, and Re means the real part. For the complex case we would analogously expect the characteristic function to be

$$\exp\{-\tfrac{1}{4}\, |z|^2 E|\xi|^2\} \qquad \forall\, z \in \mathbb{C}^1;$$

this form of characteristic function leads to the narrow definition, as equality (2) below shows. For a complex number $u = u' + iu''$ denote the pair $(u', u'') \in \mathbb{R}^2$

by $\hat{u}$. Clearly, $\mathrm{Re}\,\xi\overline{z} = (\hat{\xi}|\hat{z})$, where $(\cdot|\cdot)$ stands for the scalar product. Therefore, equality (1) gives

$$\chi_{\xi}(z) = \chi_{\hat{\xi}}(\hat{z}), \qquad \xi = \xi' + i\xi'', \qquad z = z' + iz''. \tag{2}$$

There are at least two more, nonformal, and sensible, arguments in favor of the narrow definition. They are connected with two basic properties of real Gaussian systems. It is well known that if a pair of real random variables has a jointly Gaussian distribution and the variables are uncorrelated then they are independent as well. If we want to maintain this property for the complex case we have to adopt the narrow definition. The proof of this fact is easy and we give it here for completeness. First we recall the necessary definitions.

A pair of real random variables ξ' and ξ'' is said to be *jointly Gaussian* if $\alpha'\xi' + \alpha''\xi''$ has a Gaussian distribution for any choice of real numbers α' and α'', or (equivalently) if (ξ', ξ'') is a Gaussian random vector in $\mathbb{R}^2$. In accordance with this definition, a pair of complex random variables ξ_1 and ξ_2 is said to be *jointly Gaussian* (or to constitute a *Gaussian random vector* in the two-dimensional complex Euclidean space $\mathbb{C}^2$), in the *narrow* or *wide sense* if, for any choice of complex numbers a_1 and a_2, the random variable $a_1\xi_1 + a_2\xi_2$ is Gaussian in the narrow or wide sense (respectively). Obviously, a pair of complex random variables $\xi_1 = \xi_1' + i\xi_1''$ and $\xi_2 = \xi_2' + i\xi_2''$ being jointly Gaussian in the wide sense means only that the four real random variables ξ_1', ξ_1'', ξ_2', ξ_2'' are jointly Gaussian; i.e., any real linear combination of them is a Gaussian random variable. However, ξ_1 and ξ_2, being jointly Gaussian in the narrow sense, impose a specific restriction on the distributions of ξ_1 and ξ_2. Indeed, supposing for simplicity of notation that all the four Gaussian random variables are centered (the mean values are zero), we can easily verify that the random vector (ξ_1, ξ_2) is Gaussian in the narrow sense if and only if it is Gaussian in the wide sense; both coordinates ξ_1 and ξ_2 are Gaussian in the narrow sense and the conditions

$$E\xi_1'\xi_2' = E\xi_1''\xi_2'', \qquad E\xi_1'\xi_2'' = -E\xi_1''\xi_2' \tag{3}$$

are fulfilled.

On the other hand, it is easy to see that ξ_1 and ξ_2 are independent if and only if

$$E\xi_1'\xi_2' = E\xi_1'\xi_2'' = E\xi_1''\xi_2' = E\xi_1''\xi_2'' = 0 \tag{4}$$

and they are uncorrelated (i.e., $E\xi_1\overline{\xi}_2 = 0$) if and only if

$$E\xi_1'\xi_2' = -E\xi_1''\xi_2'', \qquad E\xi_1'\xi_2'' = E\xi_1''\xi_2' \tag{5}$$

(the latter is true, of course, for general random variables as well).

Equalities (5) do not imply (4) and, hence, noncorrelatedness of coordinates does not guarantee their independence if the random vector (ξ_1, ξ_2) is Gaussian only in the wide sense. However, equalities (5), together with (3), imply (4) and, therefore, noncorrelatedness does ensure independence if the vector is Gaussian in the narrow sense.

The second argument in favor of the narrow definition comes from the 2-stability property of real Gaussian random variables (see the Remarks to Lemma 1 below). This is discussed in Section 2 where the auxiliary statements are given. The main result of the paper is proved in Section 3.

2. AUXILIARY RESULTS

LEMMA 1. *Let $\xi = \xi' + i\xi''$ be a complex random variable such that $E|\xi|^2 < \infty$, n be a natural number, $\xi_1, \xi_2, \ldots, \xi_n$ be independent copies of ξ, and $a_k = a_k' + ia_k''$ (for $1 \le k \le n$) be complex numbers. Denote $\eta = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$, $\sigma_1^2 = \mathrm{var}(\xi')$, $\sigma_2^2 = \mathrm{var}(\xi'')$ and $\rho = \mathrm{cov}(\xi', \xi'')$. If $\mathrm{var}(\eta') = \sigma_1^2$, $\mathrm{var}(\eta'') = \sigma_2^2$ and $\mathrm{cov}(\eta', \eta'') = \rho$, then*

$$|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 = 1. \tag{6}$$

Moreover, if

$$a_1''^2 + a_2''^2 + \cdots + a_n''^2 \neq 0, \tag{7}$$

then $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$.

Proof. The conditions of the lemma easily yield the following equalities:

$$\sigma_1^2 \sum_{k=1}^n a_k'^2 + \sigma_2^2 \sum_{k=1}^n a_k''^2 - 2\rho \sum_{k=1}^n a_k' a_k'' = \sigma_1^2, \tag{8}$$

$$\sigma_1^2 \sum_{k=1}^n a_k''^2 + \sigma_2^2 \sum_{k=1}^n a_k'^2 + 2\rho \sum_{k=1}^n a_k' a_k'' = \sigma_2^2, \tag{9}$$

$$(\sigma_1^2 - \sigma_2^2) \sum_{k=1}^n a_k' a_k'' + \rho \sum_{k=1}^n (a_k'^2 - a_k''^2) = \rho. \tag{10}$$

If we add equalities (8) and (9) we get equality (6). Using (6), we can rewrite (10) in the form

$$(\sigma_1^2 - \sigma_2^2) \sum_{k=1}^{n} a_k' a_k'' = 2\rho \sum_{k=1}^{n} a_k''^2. \tag{11}$$

Now suppose that conditions (7) are fulfilled. Solving (11) for ρ and substituting into (8), we get the equality

$$(\sigma_1^2 - \sigma_2^2) \left[\sum_{k=1}^{n} a_k''^2 \right]^2 = -(\sigma_1^2 - \sigma_2^2) \left[\sum_{k=1}^{n} a_k' a_k'' \right]^2$$

which shows that both sides of this equality are zero. Since (7) holds, we have $\sigma_1^2 = \sigma_2^2$; moreover, since (11) holds, we find that $\rho = 0$. ■

Remark 1. The lemma shows that Gaussian random variables in the wide sense do not have the stability property with respect to nontrivial coefficients. More precisely, if $\xi_1, \xi_2, \ldots, \xi_n$ are independent copies of a centered Gaussian random variable ξ and $a_1, a_2, \ldots, a_n$ are complex numbers (not all of which are real), $a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$cannot have the same distribution as ξ if only the wide sense definition is meant. However, it is easy to check that if ξ is Gaussian in the narrow sense then ξ has the nontrivial stability property with the index two, i.e., $a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$has the same distribution as ξ for any collection of complex numbers such that $|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 = 1$. (Note that ξ is supposed to be centered; otherwise, the additional condition $a_1 + a_2 + \cdots + a_n = 1$ is needed.)

Remark 2. In what follows, "complex Gaussian random variable" will always mean "complex Gaussian random variable in the narrow sense" unless stated otherwise. Theorem 2 below shows that this nontrivial 2-stability property characterizes the class of complex Gaussian random variables; it is not shared by any other complex random variable.

We give two more auxiliary results before formulating this theorem.

LEMMA 2. *Let ξ be a complex Gaussian random variable on some probability space (Ω, P), let χ be the characteristic function of ξ, and let $u \in \mathbb{C}^1$ be an arbitrarily given complex number. The following inequality is valid:*

$$\int_{|\xi| \le 1/|u|} [\operatorname{Re} \xi(\omega)\overline{u}]^2 P(d\omega) \le -3 \log \chi(u).$$

This lemma is an easy consequence of the corresponding statement for the real case ([5, p. 196]; see also [1, Lemma 1]). It suffices to note that the value of the characteristic function of the real random variable $\operatorname{Re} \xi\overline{u}$ at the point $t = 1$ is $\chi(u)$, and $|\operatorname{Re} \xi\overline{u}| \le 1$ if $|\xi| < 1/|u|$.

LEMMA 3. *Let f be a real-valued function, defined and continuous everywhere on the complex plane $\mathbb{C}^1$, except possibly at the point $z = 0$. Suppose f satisfies the relation*

$$f(z) = \lambda_1 f(a_1 z) + \lambda_2 f(a_2 z) + \cdots + \lambda_n f(a_n z) \qquad \text{if } z \neq 0, \tag{12}$$

where $n \geq 2$, with the positive numbers $\lambda_1, \ldots, \lambda_n$ and complex numbers $a_1, \ldots, a_n$ satisfying

$$\sum_{k=1}^{n} \lambda_k = 1$$

and

$$|a_k| < 1 \qquad (1 \leq k \leq n).$$

Then:

1. *For every nonzero point $z \in \mathbb{C}^1$, there exists a sequence $z_1, z_2, \ldots$ with $\lim_{k\to\infty} z_k = 0$ such that $f(z_k) = f(z)$ for all $k = 1, 2, \ldots$.*
2. *If the limit of f exists at $z = 0$, then f is constant.*

The proof is a slight modification of that of the analogous lemma for the real case (Lemma 2 in [1]). Clearly, it is enough to prove statement 1. Let $z \in \mathbb{C}^1$ be given. The value $f(z)$ is a convex linear combination of $f(a_1 z), f(a_2 z), \ldots, f(a_n z)$. Therefore, $f(a_j z) \leq f(z) \leq f(a_k z)$ for some pair (j, k) of indices. If the line segment joining the points $a_j z$ and $a_k z$ does not contain the point $z = 0$, then f is continuous along this segment, and so there exists a point z_1 such that $f(z_1) = f(z)$. Clearly, $|z_1| \leq \max(|a_j z|, |a_k z|) \leq \alpha|z|$, where $\alpha = \max(|a_1|, |a_2|, \ldots, |a_n|) < 1$. If this segment does contain zero, we can consider a contour consisting of a half-disk with center at zero and radius $r \leq \alpha|z|$ and two segments connecting the endpoints of this half-disk with the points $a_j z$ and $a_k z$. Obviously, on this contour there exists a point z_1 with the needed properties. Using relation (12) for the point z_1 we get (in the same manner) the point z_2 with $|z_2| \leq \alpha^2|z|$ and $f(z_2) = f(z)$. Since we can apply this argument repeatedly, the proof is complete.

3. THE MAIN RESULT

Now we are ready to prove our main result.

THEOREM 2. *Let ξ be a complex random variable, $n \geq 2$ be a natural number, $\xi_1, \xi_2, \ldots, \xi_n$ be independent copies of ξ and $a_1, a_2, \ldots, a_n$ be complex numbers such that*

$$\sum_{k=1}^{n} |a_k|^2 = 1 \tag{13}$$

and at least one of them is not a real number. If $a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$ *has the same distribution as* ξ *then* ξ *is a complex Gaussian random variable* (*see Remark* 2 *to Lemma* 1).

Proof. The hypotheses of the theorem give the following functional equation for the characteristic function χ of ξ:

$$\chi(z) = \chi(\overline{a}_1 z)\chi(\overline{a}_2 z) \ldots \chi(\overline{a}_n z). \tag{14}$$

Let us first assume that χ takes only nonnegative values. In this case clearly χ is also symmetric: $\chi(z) = \chi(-z)$ for $z \in \mathbb{C}^1$.

As in the real case, conditions (13) and (14) easily show that $\chi \neq 0$ everywhere in $\mathbb{C}^1$. Therefore, $\chi(z) > 0$ everywhere, and for any $z \in \mathbb{C}^1$ we can write

$$\chi(z) = \exp\{-\psi(z)\}, \tag{15}$$

where ψ is a continuous real-valued function of complex variable satisfying also the conditions

$$\psi(0) = 0, \qquad \psi(z) \geq 0, \qquad \psi(z) = \psi(-z), \qquad z \in \mathbb{C}^1.$$

We have to show that $\psi(z) = c|z|^2$ for some $c > 0$. Denote

$$\varphi(z) = \frac{\psi(z)}{|z|^2} \qquad \forall z \neq 0.$$

According to equalities (14) and (15) the function φ satisfies the equation

$$\varphi(z) = \sum_{k=1}^{n} |a_k|^2 \varphi(a_k z), \qquad z \neq 0.$$

By Lemma 3, for any $z \in \mathbb{C}^1$ there exists a sequence of nonzero complex numbers $z_1, z_2, \ldots$ such that $z_k \to 0$ and $\varphi(z_k) = \varphi(z)$ for all positive integers k. Now we apply Lemma 2, taking as u the numbers $z_1, z_2, \ldots$ successively. Then we obtain the inequality

$$\int [\operatorname{Re} \chi(\omega)\bar{e}_k]^2 I_k(\omega)\, dP \leq 3\varphi(z), \qquad k = 1, 2, \ldots, \tag{16}$$

where $e_k = z_k/|z_k|$, and I_k is the indicator function of the event $\{|\xi|\,|z_k| \leq 1\}$. By virtue of the compactness of the unit circle there exists a converging subsequence of the sequence (e_k), which we can assume to be (e_k) itself without loss of generality. Denote the limit by e_z. The integrand function in (16) converges to $[\operatorname{Re} \xi(\omega)\bar{e}_z]^2$ for any $\omega \in \Omega$. Therefore inequality (16) implies (by Fatou's lemma) that the function $[\operatorname{Re} \xi\bar{e}_z]^2$ is integrable. We want to prove integrability of $|\xi|^2$. It is easily seen that it suffices to show the existence of one more limit point of the sequence (e_k), linearly independent of e_z with respect to real coefficients. The easiest way to show this is the following one suggested by N. P. Kandelaki: if the distribution of ξ is invariant under the multiplication by i, then obviously along with e_z the point ie_z is also a limit point of the sequence (e_k). The random variable $\xi + i\tilde{\xi}$, where $\tilde{\xi}$ is an independent copy of ξ, has this invariance property and is symmetric (recall that the random variable ξ was provisionally supposed to be symmetric). Hence, $E|\xi + i\tilde{\xi}|^2 < \infty$. Using (e.g.) Fubini's theorem, it easily follows that $E|\xi|^2 < \infty$.

Therefore, $E|\xi|^2 < \infty$ and we can use Lemma 1 to prove that the limit of the function φ exists at zero. Actually, by our provisional assumption, we have $\chi(-z) = \chi(z)$ for all $z \in \mathbb{C}^1$, and hence, the first-order partial derivatives of ψ with respect to z' and z'' are equal to zero at zero. Moreover, by Lemma 1 the second mixed derivative of ψ at $z = 0$ is also zero, and the second derivatives with respect to z' and z'' are equal to each other at $z = 0$. We have also $\psi(0) = 0$ and thus Taylor's formula for $z = 0$ yields the following relation for a small neighborhood of zero:

$$\psi(z', z'') = \tfrac{1}{4} E|\xi|^2|z|^2 + o(|z|^2).$$

This relation shows that the limit at zero of the function φ does exist, and therefore Lemma 3 completes the proof.

Passing now to the general case, note that $|\chi|^2$ is a characteristic function of the random variable $\xi - \tilde{\xi}$ where (as above) $\tilde{\xi}$ is an independent copy of ξ. Clearly, $|\chi|^2$ is nonnegative, and it satisfies relation (14), as does ξ. Therefore, according to what has been proved already, $\xi - \tilde{\xi}$ is a (centered) Gaussian random variable. Once we apply the well-known Cramer theorem along with Lemma 1, the proof is complete. Indeed, it is easy to see that if the sum of two independent complex random variables is Gaussian either in the narrow or wide sense then both summands are Gaussian in the wide sense (but not necessarily in the narrow sense as the following simple example shows: $\xi_1 = \gamma_1 + ic\gamma_2$, $\xi_2 = c\gamma_3 + i\gamma_4$, where $c \in \mathbb{R}^1$ and γ_k are mutually independent real standard Gaussian random variables). On the other hand, if ξ is Gaussian in the wide sense and has 2-stability property it should be Gaussian in the narrow sense as well (Remark 1 to Lemma 1).

Thus we have completely proved Theorem 2. ■

Note in conclusion that in the very same manner as was made for the real case in [3], we can easily show that Theorem 2 implies the following statement. Let ξ be a Gaussian random vector in the narrow sense with values in a complex Banach space X which means that (l, ξ) is a Gaussian random variable for any continuous linear functional l in X. Then the random vector $A\xi$ is also Gaussian in the narrow sense for any measurable (not necessarily bounded) linear operator A such that the natural necessary condition $P\{\xi \in D_A\} = 1$ is satisfied.

REFERENCES

1. Vakhania, N. N. (1993), Elementary proof of Polya's characterization theorem and of the necessity of second order in the CLT, *Theory Probab. Appl.* **38**(1), 166–168.
2. Traub, J., Wasilkowski, G., and Woźniakowski, H. (1988), "Information-Based Complexity," Academic Press, New York.
3. Vakhania, N. N. (1991), Gaussian mean boundedness of densely defined linear operators, *J. Complexity* **7,** 225–231.
4. Vakhania, N. N., and Kandelaki, N. P. (1996), Random vectors with values in complex Hilbert spaces, *Theory Probab. Appl.* **41**(1), 31–52.
5. Loeve, M. (1963), "Probability Theory," Van Nostrand, Princeton.
6. Polya, G. (1923), Herleitung des Gauss'schen Fehlergesetzes aus einer Funktionalgleichung, *Math. Z.* **18,** 185–188.